Continental-scale machine learning classification of fire smoke pollution events for health research in Australia, 2001-2020

K. Gopi [a,b], J. Shao [a,b], G.G. Morgan [a,b,c], C. Yuen [d], N. Borchers-Arriagada [c,e], J. Van Buskirk [a], E. Jegasothy [a] and I.C. Hanigan [b,c,d]

[a] *Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, NSW, Australia*
[b] *Healthy Environments and Lives (HEAL), National Research Network, Australia*
[c] *Centre for Safe Air, NHMRC Centre for Research Excellence, Hobart, Tasmania, Australia*
[d] *Curtin University School of Population Health, WHO Collaborating Centre for Climate Change and Health Impact Assessment, Western Australia, Australia*
[e] *Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia*
Email: ivan.hanigan@curtin.edu.au

Abstract: Bushfire smoke is a major contributor to air pollution in Australia, yet methods to identify smoke-specific exposure remain limited, hampering epidemiological research and public health decision-making. Existing approaches relying solely on particulate matter with a diameter of 2.5 micrometres or less ($PM_{2.5}$) concentrations cannot reliably distinguish bushfire smoke from other pollution sources, creating a critical methodological gap for environmental health studies. This study aimed to develop and validate a machine learning framework to define the best combination of statistical and contextual indicators, or 'flags', for identifying landscape fire smoke days using $PM_{2.5}$ data downloaded from the Australian Bushfire Smoke Exposure Project Version 1.3 dataset (2001–2020).

The methodology involved applying classification and regression trees (CART) models trained on the previously published Australian Validated Biomass Smoke Events Database as the response variable (Hanigan et al. 2018), using statistical and satellite derived flags as predictors. The statistical indicators such as daily $PM_{2.5}$ 95th or 99th percentiles, or the remainders from seasonal-trend decomposition, were less reliable than CART models incorporating multiple predictors. The best performing CART model combined the threshold-exceedances of $PM_{2.5}$, the season (January–February and September–December), and active fires within 50 km identified from satellite data. While the model successfully identified 511 smoke days in the Validated Biomass Smoke Events Database (69.5% of actual smoke events), it showed high sensitivity to non-smoke conditions, generating 1,252 false positives — representing 71.0% of all positive predictions. Despite these limitations, the approach represents a significant methodological advance over single-indicator methods. Localisation of the CART models to specific geographic domains may improve alignment with validated events, highlighting regional variability in fire behaviour and meteorological patterns.

These findings demonstrate that combining statistical thresholds with contextual information enhances the detection of fire smoke days, supporting more accurate exposure assessment for epidemiological research. The validated framework provides the first continental-scale coverage for retrospective bushfire smoke exposure classification in Australia, enabling systematic analysis of health impacts across diverse populations and geographic regions. With increasing bushfire risk under climate change, scalable methods to identify fire-related $PM_{2.5}$ are essential to strengthen public health research and protection. This dataset and methodology provide critical research infrastructure for generating evidence to inform policy development and targeted intervention strategies.

Keywords: *wildfire smoke, event identification, exposure assessment, air quality*

1. INTRODUCTION

Bushfires are increasing globally with climate change, particularly in hot, dry regions such as Australia (Van Oldenborgh et al. 2021; Bureau of Meteorology and Commonwealth Scientific and Industrial Research Organisation 2022). Airborne particulate matter with a diameter of 2.5 micrometres or less ($PM_{2.5}$) from bushfires or controlled burns — along with other sources such as wood heaters and dust storms— is a major type of health-harming air pollution. In Australia, extreme $PM_{2.5}$ events typically arise from bushfires, hazard reduction burns, or dust storms.

Extreme $PM_{2.5}$ poses a major cardiorespiratory risk. Uncontrolled bushfires can elevate $PM_{2.5}$ for weeks to months. These fine particles penetrate deep into the lungs and enter circulation, where they trigger systemic inflammation (Franklin et al. 2015). Epidemiological studies link bushfire smoke with cardiorespiratory mortality and morbidity (Jegasothy et al. 2023), perinatal outcomes (Heft-Neal et al. 2022), and other health effects (Elser et al. 2023). However, evidence of health impacts from bushfire-specific $PM_{2.5}$ remains limited compared to urban $PM_{2.5}$ from the range of sources related to traffic and industry. As winter inversions can also elevate $PM_{2.5}$ from wood heaters (Reisen et al. 2013), bushfire smoke exposures must also be distinguished from this source, along with dust or industrial accidents such as factory fires.

In Australia, air pollution data is sparse and relies on a small number of state-run monitoring sites. For instance, New South Wales (NSW) operates 19 monitors across the Sydney Greater Metropolitan Region (GMR), covering 9801 km² and 4.9 million people in 2019 (Jegasothy et al. 2023). Few monitors exist in regional areas where bushfires often occur (Hanigan et al. 2025). Crucially, there are no monitoring observations that identify fire smoke $PM_{2.5}$ specifically, so indirect methods via modelling are utilised.

Estimates of fire-specific $PM_{2.5}$ are essential for epidemiological exposure-response analysis. To predict exposure to bushfire-specific $PM_{2.5}$, studies have used decomposition methods to statistically estimate fire smoke from $PM_{2.5}$ measured at monitors (Morgan et al. 2010), spatial interpolation (Johnston et al. 2021), chemical transport models (CTMs) (Keywood et al. 2019), hybrid and blending/ensemble approaches (Aguilera et al. 2023; Sowden et al. 2024) and machine learning (Larsen et al. 2021; Childs et al. 2022; Paul et al. 2022; Raffuse et al. 2023; Borchers-Arriagada et al. 2024). Validation of these predictions is imperative to support robust epidemiological study.

Several Australian studies have approached this task by creating the Validated Biomass Smoke Events Database (Hanigan et al. 2018). This database focusses on extreme or unusually high $PM_{2.5}$ events which are the days likely to be fire-smoke related, and uses procedures drawn from several previously published studies to validate the events. Morgan et al. (2010) defined smoke days in Sydney as PM_{10} (particulate matter with a diameter of 10 micrometres or less) above the 99th percentile plus nearby fire activity, estimating bushfire PM_{10} as the excess over surrounding days. Johnston et al. (2011) utilised expert review of news archives, government reports, and satellite imagery. Jegasothy et al. (2023) used the 95th percentile for $PM_{2.5}$, confirmed with satellite imagery, and applied a statistical method known as seasonal-trend decomposition using loess (STL) to separate partition out smoke $PM_{2.5}$ from total $PM_{2.5}$.

2. METHODS

This study used machine learning to define the best combination of predictors (or 'flags') for identifying landscape fire smoke days in the Bushfire Smoke Exposure Project's $PM_{2.5}$ Model Version 1.3 (Hanigan et al. 2023; Borchers-Arriagada et al. 2024). This is a continental scale Australian modelled dataset for daily $PM_{2.5}$ predictions from 2001 to 2020 with 5×5 km^2 spatial resolution. We trained a classification algorithm using smoke events from the Validated Biomass Smoke Events Database to improve the accuracy of $PM_{2.5}$ attribution.

2.1. Australian Bushfire Smoke PM2.5 - V1.3 Model Prediction Database

In our previous work we estimated daily $PM_{2.5}$ at 5×5 km^2 resolution across Australia (2001–2020) using a random forest model with a wide range of remotely sensed data (eg: active fires, AOD), land-use, climatic, meteorological, emission source and topographic predictors (Hanigan et al. 2023; Borchers-Arriagada et al. 2024). Briefly, the model was trained using monitor data, satellites, weather, climate variables and land use as predictors and then STL decomposition produced seasonal, trend, and remainder components (Cleveland et al. 1990). Full details are available online at https://cardat.github.io/bushfire_pm25_v1_3_tools.

Extreme days (i.e. outliers) were defined for each pixel in two ways: first if the remainder on the day exceeded two standard deviations (SD) of all the remainders in the time-series for that pixel (where remainder is the difference between the prediction on that day and the season-and-trend estimate on that day). Second if the prediction was greater than the 95th percentile. Because the protracted extreme bushfire smoke events in 2019-2020 identified multiple outliers that made the STL decomposition unreliable, we first trimmed the remainder time-series to the 99th percentile before calculating statistical thresholds. Flags from satellite hot spots, dust estimates, and temperature were also considered to help identify event days. Protocols were developed to flag dust, wood, and fire smoke days.

2.2. Validated Biomass Smoke Events Database

A Biomass Smoke Validated Events Database of extreme fire smoke events has previously been developed using news, agency records, and satellites (Hanigan et al. 2018) for selected locations and time periods within Australia and we define "validated landscape fire event days" as those identified in the database. The Biomass Smoke Validated Events Database includes four related protocols for validation of events in observational datasets: **Morgan 2010:** 99th percentile and newspaper archives (Morgan et al. 2010); **Johnston 2011:** 95th percentile and multiple records/dust database (Mctainsh 1998; Johnston et al. 2011); **Salimi 2016/2017:** satellite, news, and government reports (Salimi et al. 2017); and **Bare Minimum:** any contributor-evidence (Williamson et al. 2016; Kjellstrom and Hanigan 2004).

2.3. Flag Development and Thresholds

This project flagged extreme $PM_{2.5}$ days using thresholds and external data such as The Moderate Resolution Imaging Spectroradiometer (MODIS) active fire data, aerosol optical depth estimates from the Copernicus Atmosphere Monitoring Service (CAMS), dust surface mass concentration from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2), and surface temperature data from the Australian Bureau of Meteorology Gridded Climate Data. One key flag identified days when STL remainder exceeded two standard deviation (SD) of the trimmed remainder series. A second threshold-based flag identified days where total $PM_{2.5}$ exceeded the 95th percentile.

2.4. Machine learning algorithm for classification

The machine learning method of Classification and Regression Tree (CART) was used (Breiman 2001). This is an earlier form of the regularly used Random Forest algorithm (Breiman 2001) and is routinely derived to build decision rules based on recursive partitioning of data by iteratively splitting the data to maximise between-group deviance and minimise within-group deviance. We applied the CART model to predict validated smoke events (from the Biomass Smoke Validated Events Database) as the response variable and used our statistical flags or satellite data as predictors. The trained CART models allow prediction of smoke days across the national gridded $PM_{2.5}$ dataset even where validated fire smoke events data is not available in the database.

3. RESULTS

The Validated Smoke Events Database covers 16 study sites with varying periods (Figure 1). The study locations used in the Biomass Smoke Validated Events Database 2001–2020 were: Albany, Albury, Bunbury, Busselton, Bathurst, Geraldton, Illawarra, Melbourne, Newcastle, Perth, Sydney, Tamworth, Wagga Wagga, Hobart and Launceston.

Figure 1: Map of study sites covered by the Extended Biomass Validated Smoke Events Database 2001 to 2020

3.1. Comparisons with Bushfire-specific smoke PM2.5 V1.3 estimates and flags with the Validated Events Database

Predicted daily $PM_{2.5}$ values were averaged across 5 × 5 km^2 cells. To avoid skew from zeros, predictions were truncated at the 25^{th} percentile. Initial comparisons showed that the statistical flags alone were unreliable. Some sites aligned reasonably with validated events, but many smoke days fell below thresholds, while others were falsely flagged (see online supplementary material for details: OSF site “cardat_bushfire_pm25_v1_3_flags_paper_supplement” at https://osf.io/avfng).

3.2. Classification tree results

The initial CART included all flags, selecting $PM_{2.5}$ >95^{th} percentile, remainder $PM_{2.5}$ >2 SD, mean air temperature ≥12–15°C, and active fires within 5–100 km. After pruning, the CART model retained $PM_{2.5}$ >95^{th} percentile, temperature ≥15°C, fires within 50 km, and $PM_{2.5}$ >2 SD. Compared with validated data, it correctly classified 489 smoke days but produced 1274 false positives and 264 false negatives Adding “month” improved accuracy: 511 true positives, 1252 false positives, and 224 false negatives. Including a ±1-day lag increased true positives (725) but also raised false negatives (420) (Table 1). The optimal combination was $PM_{2.5}$ >95^{th} percentile, month, and fires within 50 km (Figure 2).

Table 1. Predictions from the CART model using the flags and statistical thresholds provided with the $PM_{2.5}$ *Version 1.3 dataset.*

		Classification and Regression Tree prediction	
		Not a landscape fire smoke event	Landscape fire smoke event
True value of outcome variable	Not a landscape fire smoke event	38428	1252
	Landscape fire smoke event	224	511

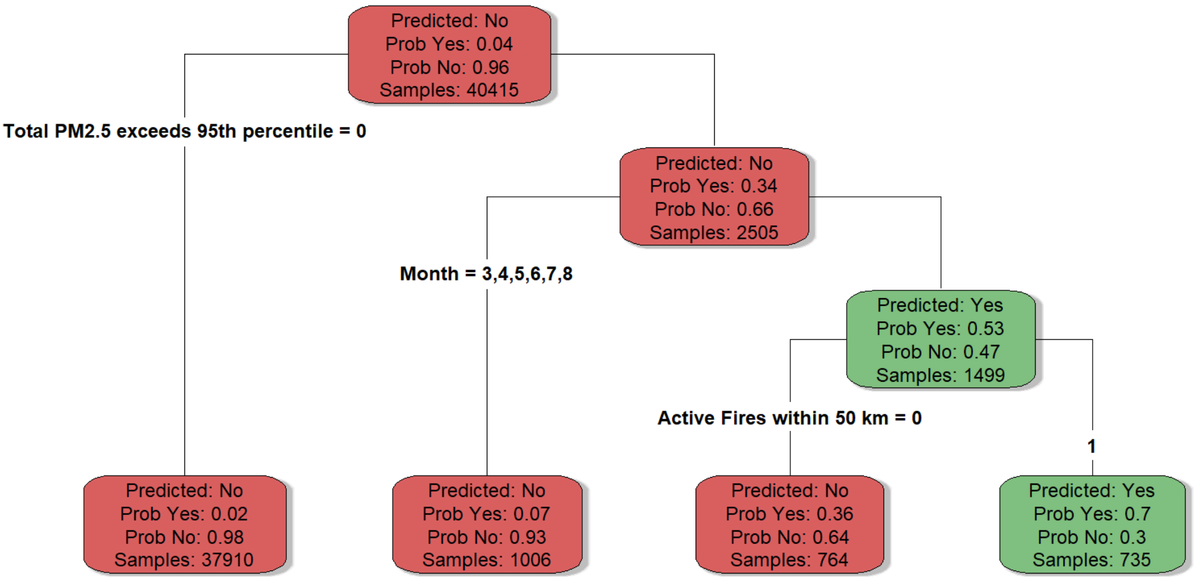

Figure 2. CART model predicting bushfire smoke days using levels, month, and fire activity. Tree splits first on $PM_{2.5}$ *exceeding 95th percentile, then by month and fires within 50km. Red nodes = No, green nodes = Yes.*

3.3. Assessing the influence of location on model accuracy

To determine whether the location influenced model accuracy, we tested the ability of the best performing CART model to predict fire smoke affected days and compared this to the validated events database. Newcastle in NSW was selected as the study region for this test (Figure 3). The CART model trained on all the study regions had the best performance from our previous tests and predicted 66 days as being affected by bushfire smoke.

4. DISCUSSION AND CONCLUSION

This study developed and validated a machine learning framework using CART to predict landscape fire smoke days across Australia from 2001–2020, achieving reliable discrimination between smoke-affected and non-smoke days through the combination of statistical outlier detection on $PM_{2.5}$ concentrations (e.g. 2SD over the remainder from STL and/or >95th percentile), seasonal timing (January–February or September–December), and local satellite observed fire activity (within 50 km). The validated model provides the first nationally consistent method for retrospectively identifying bushfire smoke exposure days in the gridded dataset, filling a critical gap in environmental health research where accurate exposure classification has been a perennial challenge.

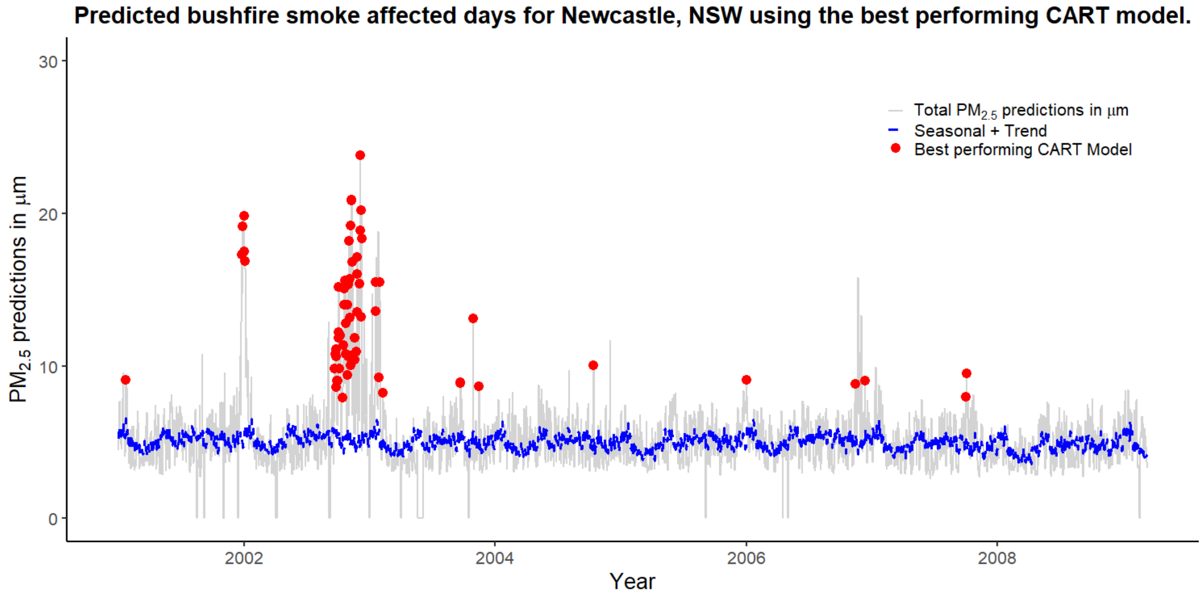

Figure 3: Number of bushfire smoke days in Newcastle, Australia according to the best performing CART model.

This work represents a significant methodological advance in bushfire smoke exposure assessment by moving beyond reliance on $PM_{2.5}$ concentrations and statistical thresholds alone. Our findings confirm earlier observations by Johnston et al. (2011) that daily $PM_{2.5}$ measurements cannot reliably distinguish smoke from other pollution sources, but demonstrate that incorporating satellite observed fire activity and seasonal context substantially improves classification accuracy. This aligns with work by Jegasothy et al. (2023) emphasising the importance of a multi-indicator approach.

The integration of satellite fire data with ground-based $PM_{2.5}$ measurements builds upon established approaches while offering improved precision compared to previous studies (Cortes-Ramirez et al. 2022). Unlike earlier methods that focused on extreme smoke events (Jegasothy et al. 2023; Hertzog et al. 2024), our framework captures a more complete spectrum of smoke-affected days, providing a more comprehensive foundation for epidemiological research. The use of the expert-reviewed Validated Smoke Event Database records (Hanigan et al. 2018) to train the CART model represents a methodological innovation beyond the work conducted by prior studies.

The primary strength of this study lies in utilising validated smoke event records compiled from multiple sources including satellite data, news articles, government reports, and meteorological assessments. The national scope and 20-year temporal coverage provide comprehensive scale for bushfire smoke exposure research in Australia and lend support for generalisability across geographic and temporal study domains. The CART methodology offers a robust, widely tested machine learning framework.

Several limitations are relevant to application of these findings. The Validated Biomass Smoke Event Database relies on expert judgment applied to available observational data, and may itself contain uncertainty in smoke day identification, particularly for less severe events that are less visible to remote sensing or generate limited media attention or government reports. Moreover, our optimal model still produced substantial misclassification, missing 224 genuine smoke-affected days while incorrectly flagging 1,252 non-fire days. This error rate reflects the inherent challenge of distinguishing smoke from other $PM_{2.5}$ sources using available data. Furthermore, our analysis revealed significant regional variability that suggests a single national model may not capture local fire behaviour and meteorological patterns effectively, consistent with findings by de Jesus et al. (2020) regarding spatial heterogeneity in air pollution patterns. Finally, there are several other pollutants that are emitted during a bushfire event such as black carbon, ash, volatile organic compounds, and nitrogen oxides. These pollutants are also harmful to human health but estimating concentrations of smoke related pollutants beyond $PM_{2.5}$ was outside the scope of this project.

The application of this methodology is particularly timely given the evidence of increasing fire activity under climate change (Van Oldenborgh et al. 2021). Public health agencies can utilise this framework to identify historical patterns of smoke exposure for vulnerable communities, informing targeted intervention strategies and resource allocation. However, users must exercise caution in applying these classifications in their specific studies, particularly recognising that the method is geared toward capturing days with elevated or extreme landscape fire smoke, rather than all days affected by fire smoke. The substantial false positive rate suggests that additional verification may be necessary for studies requiring high specificity in exposure classification.

Future research should incorporate the recently released ARDC Historical Bushfire Boundaries dataset (Ryu and Charalambou 2023). This presents an opportunity for independent validation and potential model refinement using more precise fire perimeter data. Additionally, investigating optimal temporal windows for exposure classification remains important, as our lag analysis suggested that smoke exposure effects may span multiple days, though incorporating temporal lags increased false positive rates in the current framework.

In conclusion, this study provides a validated framework for classifying smoke-affected days using statistical and satellite-based observations of fire activity indicators. A national CART model offers a foundation, but local refinement may be needed. The findings of this work has significant implications for public health. Fire-related $PM_{2.5}$ is linked to respiratory, cardiovascular, and perinatal outcomes. Yet health studies lack accurate bushfire-specific exposure estimates to clarify these associations. This dataset enables improved nationwide research. With climate change driving more frequent bushfires, scalable and precise smoke exposure methods are critical for public health research and protection.

5. ACKNOWLEDGEMENT

The Bushfire Smoke Exposure Project received funding support from the Australian Research Data Commons (ARDC) Bushfire Data Challenges project (https://doi.org/10.47486/DC008) as well as a small seed grant from the Centre for Safe Air (CSA; https://safeair.org.au/) which is funded by the National Health and Medical Research Council (NHMRC Grant No. 2015584) and the HEAL Network (NHMRC Grant No. 2008937). The resources of the Clean Air Research Data and Analysis Tools (CARDAT) platform (https://cardat.github.io) were used. CARDAT is supported by CSA and the WHO Collaborating Centre for Climate Change and HIA. The Claude AI helped refine language. All content has been verified.

6. REFERENCES

Aguilera, Rosana, Nana Luo, Rupa Basu, et al. 2023. 'A Novel Ensemble-Based Statistical Approach to Estimate Daily Wildfire-Specific PM2.5 in California (2006–2020)'. *Environment International* 171 (January): 107719. https://doi.org/10.1016/j.envint.2022.107719.

Borchers-Arriagada, Nicolas, Geoffrey G Morgan, Joseph Van Buskirk, et al. 2024. 'Daily PM2.5 and Seasonal-Trend Decomposition to Identify Extreme Air Pollution Events from 2001 to 2020 for Continental Australia Using a Random Forest Model'. *Atmosphere* 15 (1341). https://doi.org/10.3390/atmos15111341.

Breiman, L. 2001. 'Random Forests'. *Machine Learning* 45 (1): 5–32. https://doi.org/10.1023/a:1010933404324.

Bureau of Meteorology and Commonwealth Scientific and Industrial Research Organisation. 2022. 'State of the Climate 2022'. https://www.csiro.au/en/research/environmental-impacts/climate-change/State-of-the-Climate.

Childs, M. L., J. Li, J. Wen, and others. 2022. 'Daily Local-Level Estimates of Ambient Wildfire Smoke PM2.5 for the Contiguous US'. *Environmental Science & Technology* 56: 13607–21.

Cleveland, R. B., W. S. Cleveland, J. E. Mcrae, and I. Terpenning. 1990. 'STL: A Seasonal-Trend Decomposition'. *J. Off. Stat* 6: 3–73.

Cortes-Ramirez, J., R.N. Michael, L.D. Knibbs, H. Bambrick, M.R. Haswell, and D. Wraith. 2022. 'The Association of Wildfire Air Pollution with COVID-19 Incidence in New South Wales, Australia'. *Science of The Total Environment* 809 (February): 151158. https://doi.org/10.1016/j.scitotenv.2021.151158.

Elser, H., S. T. Rowland, M. S. Marek, and others. 2023. 'Wildfire Smoke Exposure and Emergency Department Visits for Headache: A Case-Crossover Analysis in California, 2006-2020'. *Headache* 63: 94–103.

Franklin, Barry A, Robert Brook, and C. Arden Pope. 2015. 'Air Pollution and Cardiovascular Disease'. *Current Problems in Cardiology* 40 (5): 207–38. https://doi.org/10.1016/j.cpcardiol.2015.01.003.

Hanigan, IC, GG Morgan, GJ Williamson, et al. 2018. 'Extensible Database of Validated Biomass Smoke Events for Health Research'. *Fire* 1 (3): 1–8. https://doi.org/10.3390/fire1030050.

Hanigan, I.C., W. Yu, C. Yuen, et al. 2025. 'Deep Ensemble Machine Learning with Bayesian Blending Improved Accuracy and Precision of Modelled Ground-Level Ozone for Region with Sparse Monitoring: Australia, 2005–2018'. *Environmental Modelling & Software* 187 (April): 106378. https://doi.org/10.1016/j.envsoft.2025.106378.

Hanigan, Ivan, Cassandra Yuen, Karthik Gopi, Nicolas Borchers-Arriagada, Joe Van Buskirk, and Geoffrey Morgan. 2023. 'Bushfire_specific_PM25_Aus_2001_2020_v1_3'. Version 1.3. CAR. https://cloud.cardat.org/index.php/apps/files/?dir=/ResearchProjects_CAR/Bushfire_Smoke_PM_Exposures/Bushfire_specific_PM25_Aus_2001_2020_v1_3.

Heft-Neal, S., A. Driscoll, W. Yang, and others. 2022. 'Associations between Wildfire Smoke Exposure during Pregnancy and Risk of Preterm Birth in California'. *Environmental Research* 203: 111872.

Hertzog, Lucas, Geoffrey G. Morgan, Cassandra Yuen, et al. 2024. 'Mortality Burden Attributable to Exceptional PM2.5 Air Pollution Events in Australian Cities: A Health Impact Assessment'. *Heliyon* 10 (2): e24532. https://doi.org/10.1016/j.heliyon.2024.e24532.

Jegasothy, Edward, Ivan C. Hanigan, Joe Van Buskirk, et al. 2023. 'Acute Health Effects of Bushfire Smoke on Mortality in Sydney, Australia'. *Environment International* 171 (January): 107684. https://doi.org/10.1016/j.envint.2022.107684.

Jesus, AL de, H Thompson, LD Knibbs, et al. 2020. 'Two Decades of Trends in Urban Particulate Matter Concentrations across Australia'. *Environmental Research* 190 (November). https://doi.org/10.1016/j.envres.2020.110021.

Johnston, Fay H., Nicolas Borchers-Arriagada, Geoffrey G. Morgan, et al. 2021. 'Unprecedented Health Costs of Smoke-Related PM2.5 from the 2019–20 Australian Megafires'. *Nature Sustainability* 4 (1): 42–47. https://doi.org/10.1038/s41893-020-00610-5.

Johnston, F.H., I.C. Hanigan, S.B. Henderson, et al. 2011. 'Creating an Integrated Historical Record of Extreme Particulate Air Pollution Events in Australian Cities from 1994 to 2007'. *Journal of the Air and Waste Management Association* 61 (4). https://doi.org/10.3155/1047-3289.61.4.390.

Keywood, Melita, Paul Selleck, Fabienne Reisen, et al. 2019. 'Comprehensive Aerosol and Gas Data Set from the Sydney Particle Study'. *Earth System Science Data* 11: 1883–903. https://doi.org/10.5194/essd-11-1883-2019.

Kjellstrom, T. K., and I Hanigan. 2004. *Health Impacts on Canberra Residents of Smoke from the January 2003 Bushfires: Report to ACT Health*. ACT Government.

Larsen, A, I Hanigan, BJ Reich, et al. 2021. 'A Deep Learning Approach to Identify Smoke Plumes in Satellite Imagery in Near-Real Time for Health Risk Communication'. *Journal of Exposure Science and Environmental Epidemiology* 31 (1): 170–76. https://doi.org/10.1038/s41370-020-0246-y.

Mctainsh, G. H. 1998. 'Dust Storm Index'. In *Sustainable Agriculture: Assessing Australia's Recent Performance: A Report to Standing Committee on Agriculture and Resource Management (SCARM) of the National Collaborative Project on Indicators for Sustainable Agriculture*. National Collaborative Project on Indicators for Sustainable Agriculture.

Morawska, L, T Zhu, N Liu, et al. 2021. 'The State of Science on Severe Air Pollution Episodes: Quantitative and Qualitative Analysis'. *Environment International* 156 (November). https://doi.org/10.1016/j.envint.2021.106732.

Morgan, Geoffrey, Vicky Sheppeard, Behnoosh Khalaj, et al. 2010. 'Effects of Bushfire Smoke on Daily Mortality and Hospital Admissions in Sydney, Australia.' *Epidemiology (Cambridge, Mass.)* 21 (1): 47–55. https://doi.org/10.1097/EDE.0b013e3181c15d5a.

Paul, Naman, Jiayun Yao, Kathleen E. McLean, David M. Stieb, and Sarah B. Henderson. 2022. 'The Canadian Optimized Statistical Smoke Exposure Model (CanOSSEM): A Machine Learning Approach to Estimate National Daily Fine Particulate Matter (PM2.5) Exposure'. *Science of The Total Environment* 850 (December): 157956. https://doi.org/10.1016/j.scitotenv.2022.157956.

Raffuse, Sean, Susan O'Neill, and Rebecca Schmidt. 2023. 'A Model for Rapid Wildfire Smoke Exposure Estimates Using Routinely-Available Data - Rapidfire v0.1.3'. Preprint, July 6. https://doi.org/10.5194/egusphere-2023-1031.

Reisen, F., C. P. Meyer, and M. D. Keywood. 2013. 'Impact of Biomass Burning Sources on Seasonal Aerosol Air Quality'. *Atmospheric Environment* 67: 437–47.

Ryu, G., and C. Charalambou. 2023. 'Historical Bushfire Boundaries'. Geoscience Australia. https://doi.org/10.26186/147763.

Salimi, Farhad, Sarah B. Henderson, Geoffrey G. Morgan, Bin Jalaludin, and Fay H. Johnston. 2017. 'Ambient Particulate Matter, Landscape Fire Smoke, and Emergency Ambulance Dispatches in Sydney, Australia'. *Environment International* 99: 208–12. https://doi.org/10.1016/j.envint.2016.11.018.

Sowden, M, IC Hanigan, DJV Robbins, M Cope, JD Silver, and J Noonan. 2024. 'Characterizing Smoke Haze Events in Australia Using a Hybrid Approach of Satellite-Based Aerosol Optical Depth and Chemical Transport Modeling'. *Remote Sensing* 16 (7). https://doi.org/10.3390/rs16071266.

Van Oldenborgh, G. J., F. Krikken, S. Lewis, and others. 2021. 'Attribution of the Australian Bushfire Risk to Anthropogenic Climate Change'. *Nat. Hazards Earth Syst. Sci.* 21: 941–60.

Williamson, Grant J, D M J S Bowman, O F Price, S B Henderson, and F H Johnston. 2016. 'A Transdisciplinary Approach to Understanding the Health Effects of Wildfire and Prescribed Fire Smoke Regimes'. *Environmental Research Letters* 11 (12): 125009. https://doi.org/10.1088/1748-9326/11/12/125009.

